A Search For The Higgs Boson In H→WW

SUSY11: Supersymmetry 2011, 28 Aug–2 Sep 2011, Fermilab

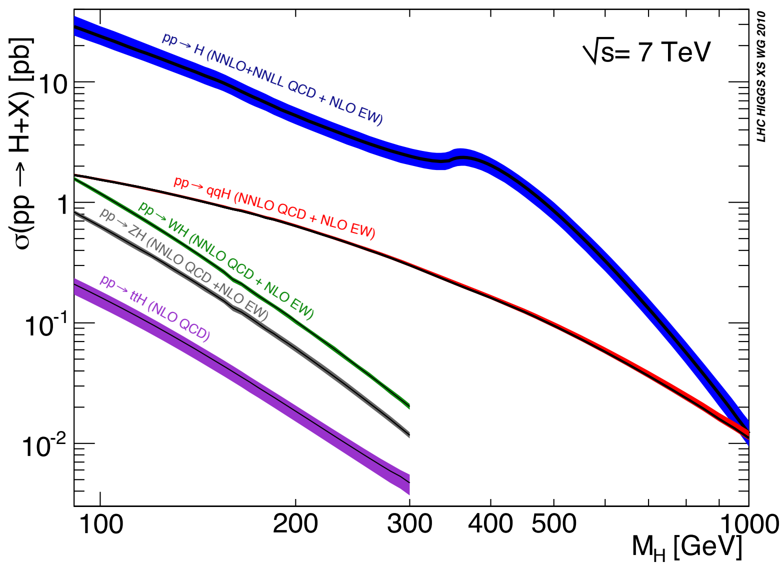
G.Cerati (UCSD)

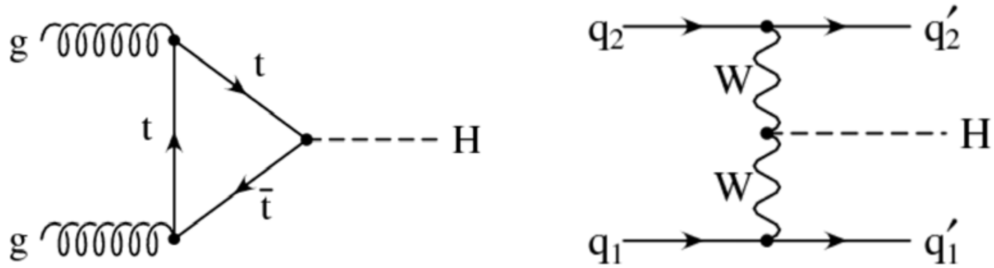
on behalf of the CMS Collaboration

L=36/pb: Phys. Lett. B699 (2011) 25–47
L=1.1/fb: CMS-PAS-HIG-11-003
L=1.5/fb: CMS-PAS-HIG-11-014

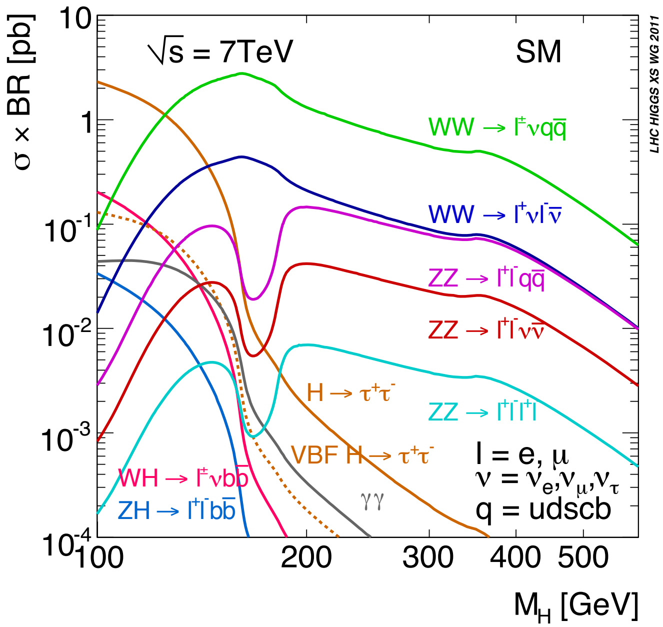
Introduction

Introduction



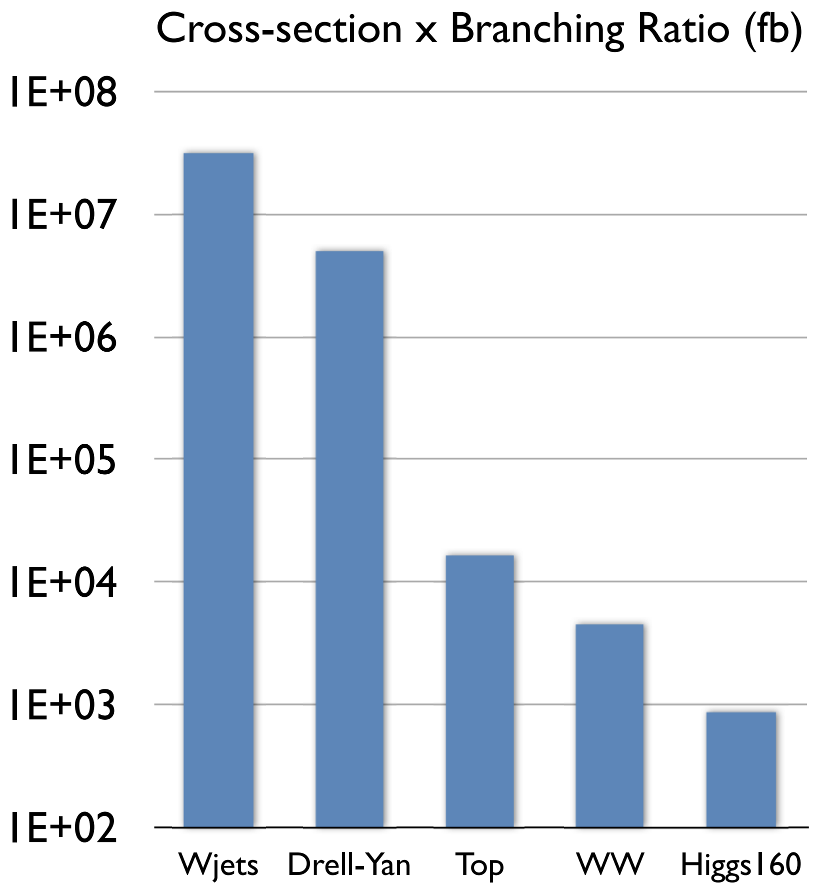


SM Higgs produced mainly via gg fusion and secondarily via VBF (→H+2j)

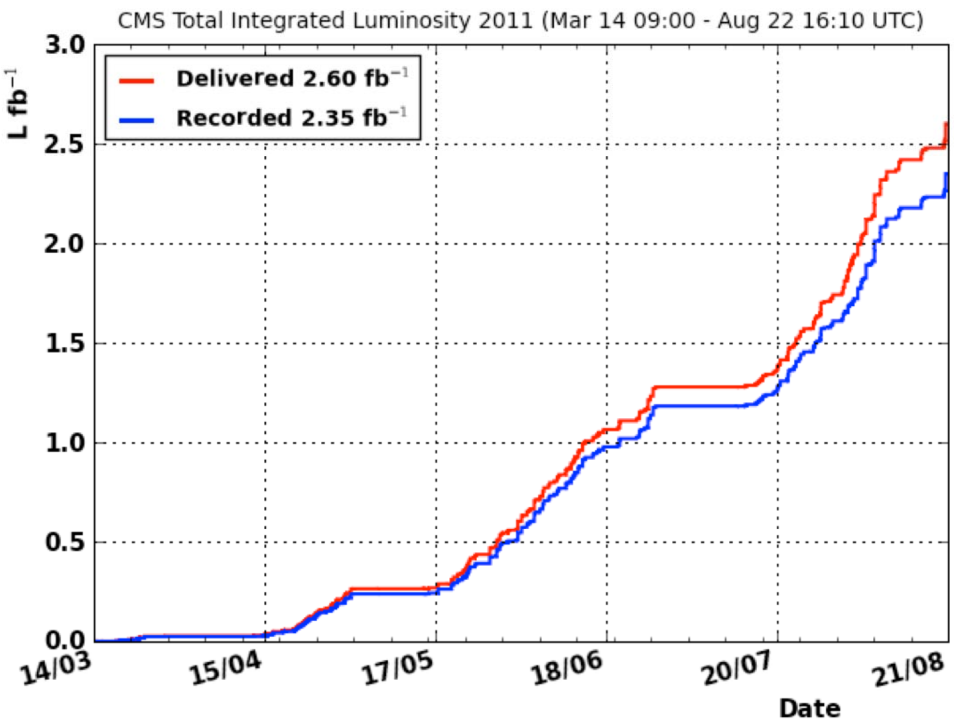


The Higgs search in the **WW→2l2ν** mode is very sensitive for a wide range of Higgs masses

Experimental Challenge



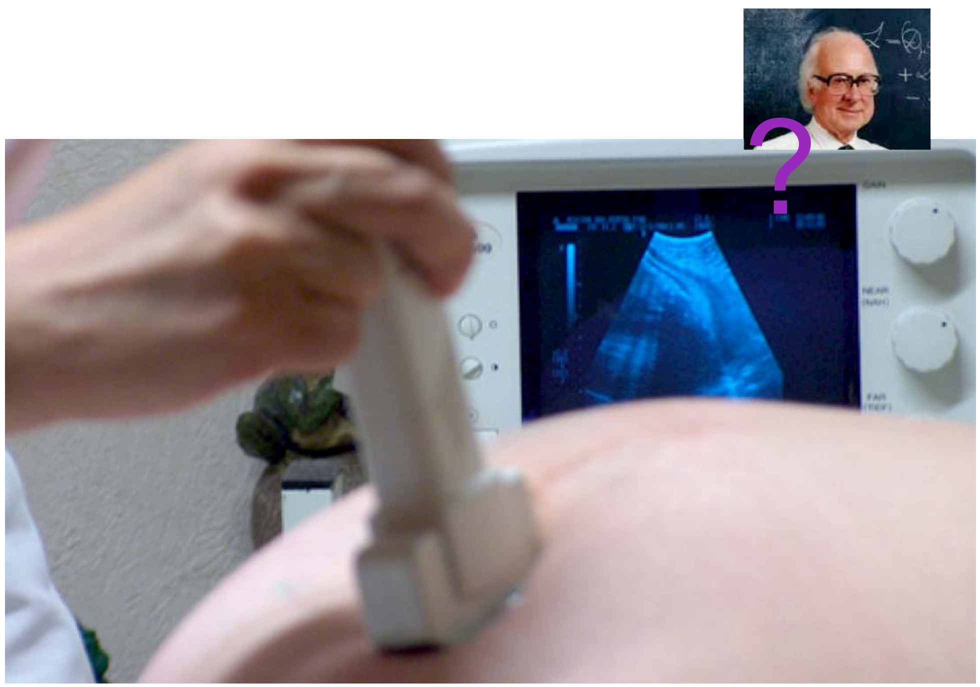
Backgrounds overwhelm the Higgs signal by several orders of magnitude



Dramatic increase in instantaneous luminosity: trigger and pile-up issues

The Higgs search is an extreme experimental challenge!

Analysis Methods

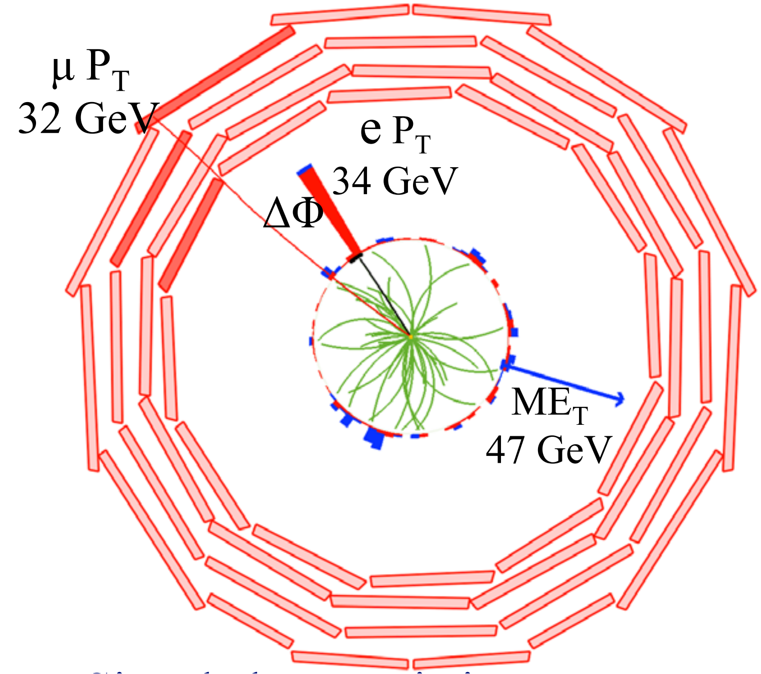


Analysis strategy summary

Final state with 2 isolated leptons
and large missing energy.

Spin correlation leads to moderately small
opening angle between the leptons.

Counting experiment, no mass peak.



Preselection:

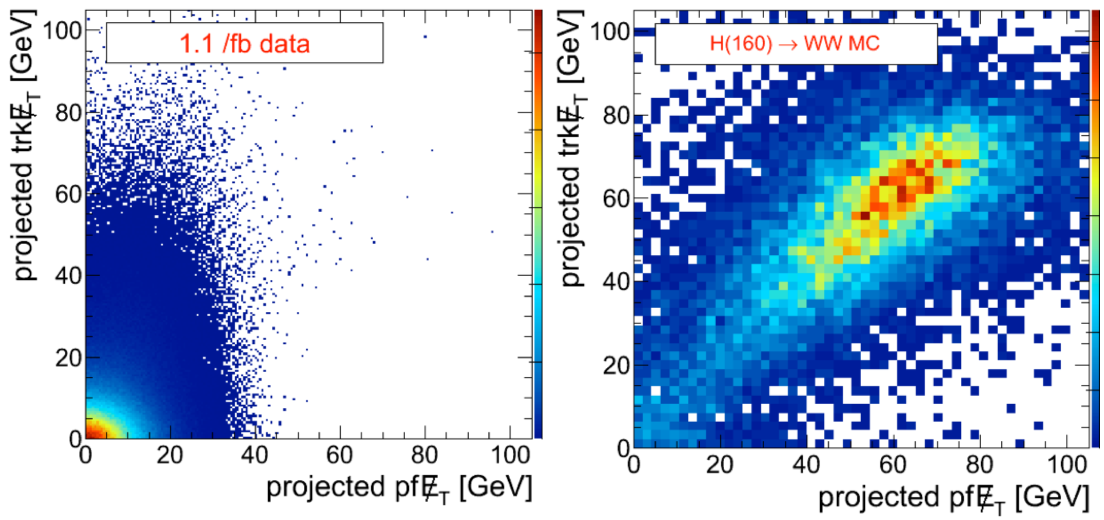
- two opposite charged, moderate p_T leptons (p_T>20,**10** GeV/c)
  - **maximize sensitivity at low m_H**
- tight identification and isolation
  - Wjets and QCD
- large missing energy (MET) and Z-mass veto, OF/SF categorization
  - Drell-Yan
- number of jets categorization (p_T>30 GeV, 0/1/2 jet bins), b-tagging veto
  - top

Higgs selection:

- mass-dependent kinematic cuts (Δφ, m_{ll}, m_T, lepton p_T)
  - WW

Dealing with pile-up

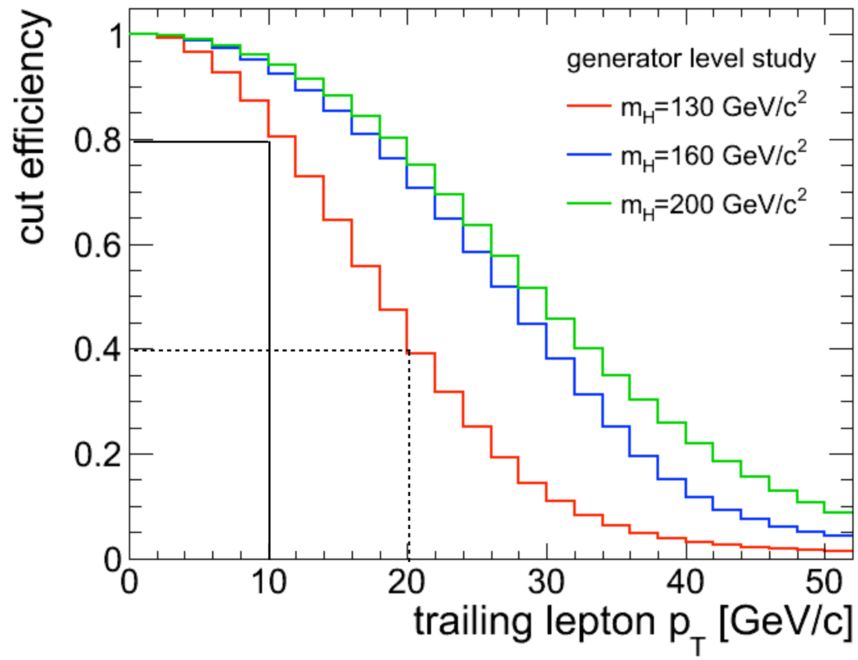
- High pile-up conditions can severely affect the precision of several **measurements** and needs dedicated **cures**
- **Vertex separation**:
  - improved reconstruction algorithm to **distinguish vertices with Δz>1 mm**;
- **Lepton isolation and identification**:
  - use only **pile-up safe variables** for identification;
  - **particle candidate-based isolation** definition less sensitive to pile-up;
- **Jet Counting**:
  - **correct the jet energy** subtracting the event average pile-up contribution;
- **Missing energy**:
  - use projected MET to reduce DY→ττ and lepton mis-measurements
  - **new MET definition** made of charged candidates from the primary vertex;
  - cut on **min between new and standard MET**: PU-safe, high rejection power.



Low p_T leptons

- A **high p_T cut** on the trailing lepton **reduces the signal efficiency** for low mass Higgs;

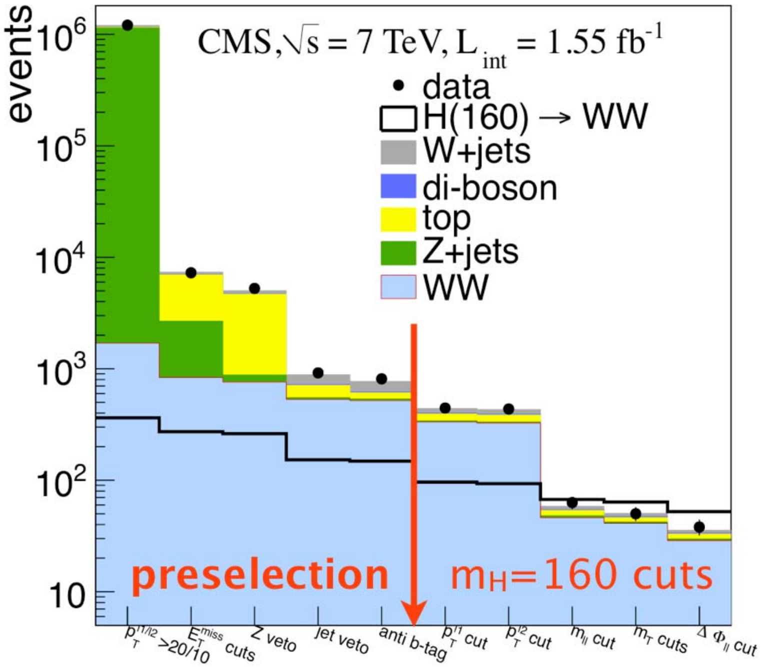
- On the other hand, **fake leptons** have a **steeply falling p_T** spectrum, and a low cut would lead to a **huge** increase of W+jets and QCD **backgrounds**;



- Optimize the lepton selection by **lowering the minimum p_T** of the trailing lepton **down to 10 GeV** and gradually **tightening the selection** in the fake-dominated regions (low pT, large η)
  - keep low fake-induced background
  - recover at least part of the signal efficiency at low p_T

Background Estimation

Backgrounds after preselection

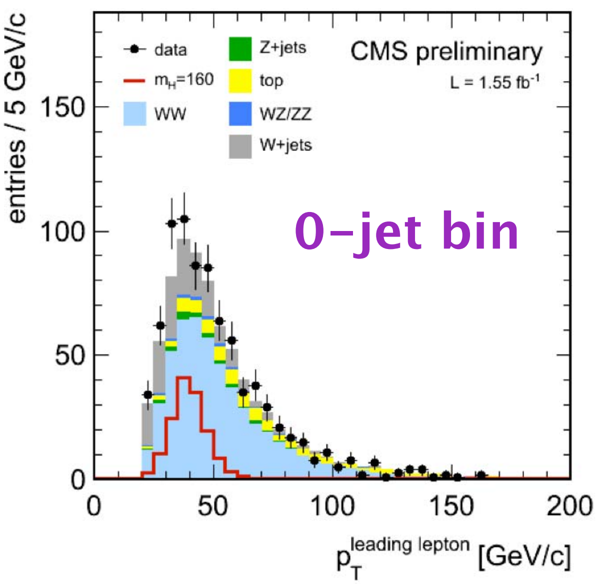


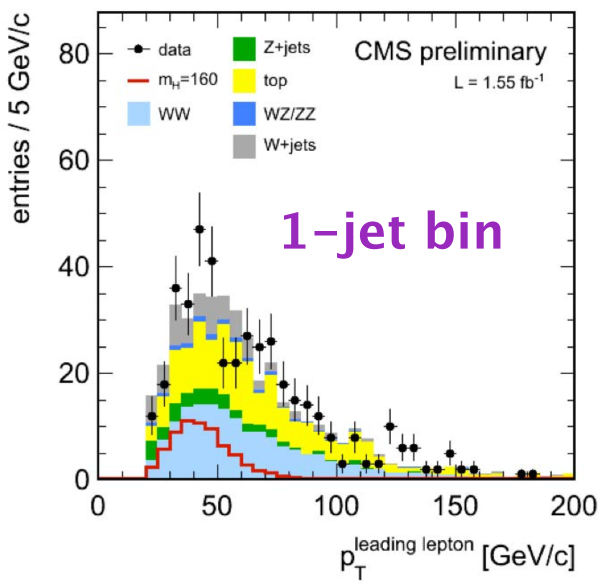
After preselection cuts:

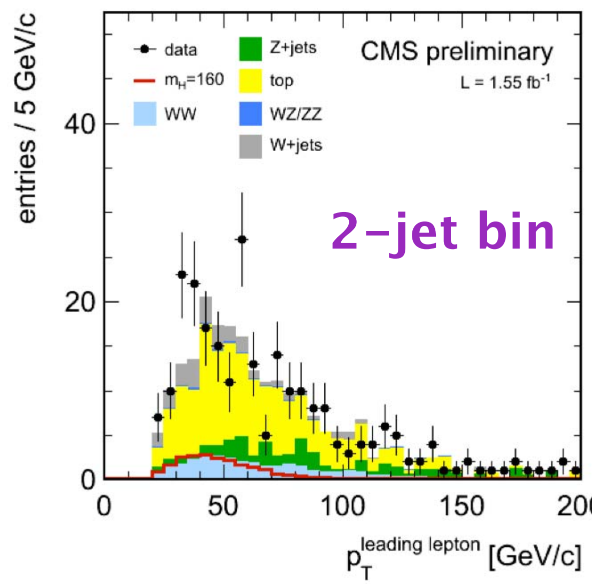
the 0-jet bin is dominated by **WW** and **W+jets**

the 1-jet bin is dominated by **WW** and **Top**

the 2-jet bin is dominated by **Top**





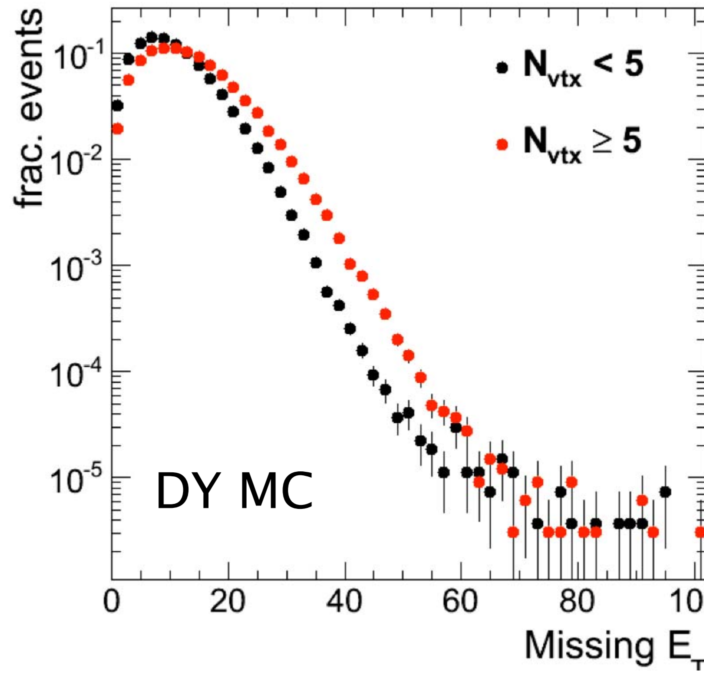


Drell-Yan Background

The Drell–Yan background is highly suppressed by a tight MET requirement.

MET resolution is affected by pile–up and the MET distribution is poorly described in simulation.

Need to **normalize** the number of Drell–Yan events in data **after MET–related selections**: $R_{out/in}$ method



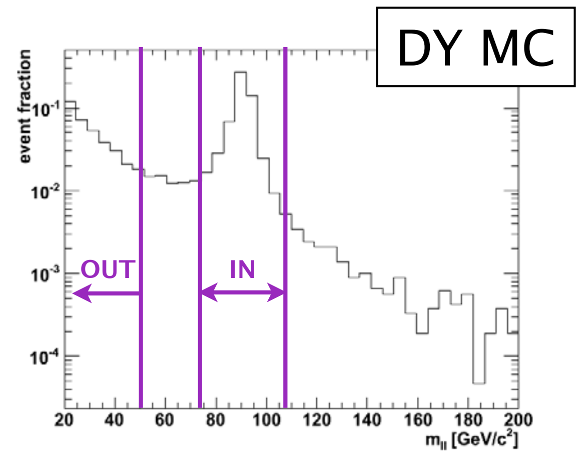
Z–peak to signal region ratio from MC, verified in data

opposite–flavor events in Z peak

$$N_{\text{out}}^{\ell\ell,\text{exp}} = \overline{R_{\text{out/in}}^{\ell\ell}}(\underline{N_{\text{in}}^{\ell\ell}} - \overline{N_{\text{in}}^{\text{non}-Z}} - \underline{N_{\text{in}}^{ZV}})$$

same–flavor events in Z peak

WZ, ZZ contribution from MC



DY estimate has **large uncertainties**:
Large systematic uncertainty due to R determination: largest deviation of the R vs MET
Large statistical uncertainty due to very few events in the Z window in data after all other selections

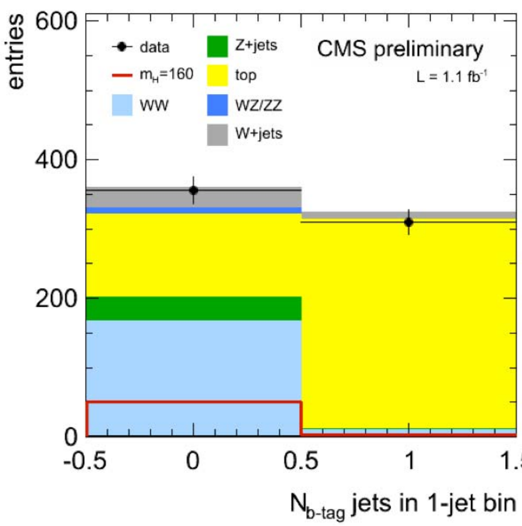
Top and Fake-induced background

Top and fake-induced background are estimated with similar methods:
invert one cut, measure cut efficiency in control region,
get residual contamination from the formula

$$N_{pass} = N_{fail} \cdot (1-\varepsilon_{fail})/\varepsilon_{fail}$$

Top background

Based on b-tagging.
Use different cuts and control regions for 0/1/2 jets.
Not enough events in signal box: normalize at preselection level and scale to signal region from MC.



Fake-induced background

method based on the measurement of the efficiency for a fake lepton to pass the analysis cuts (“fake rate”);

fake rate used to extract the number of background events passing all cuts (from the sample where one of the two leptons fails the tight ID, all other event selections are passed);

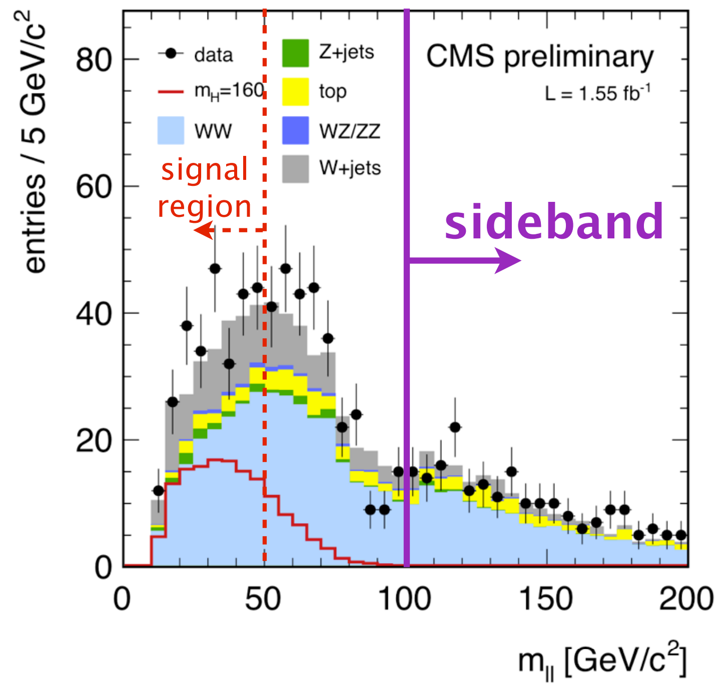
closure tests on MC and on same-sign events show very good agreement.

WW background

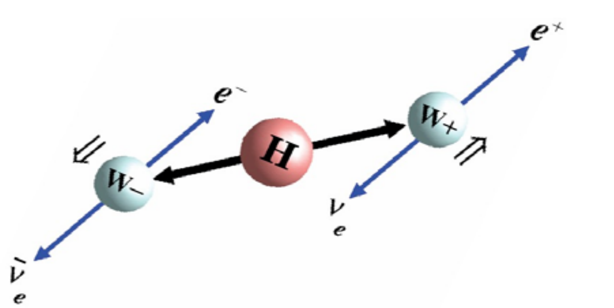
EWK WW production is the largest, irreducible background after pre-selection.

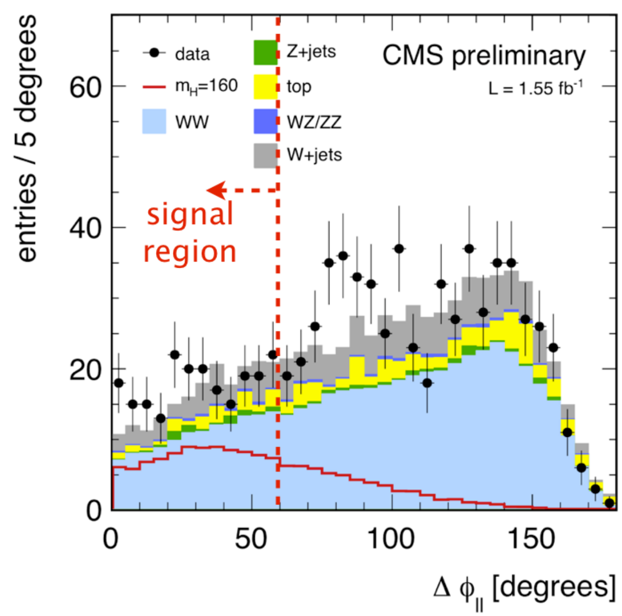
For low mass Higgs ($m_H \leq 200$ GeV/c^2), the sideband with $m_{ll} > 100$ GeV is signal-free → WW background calibrated in this region

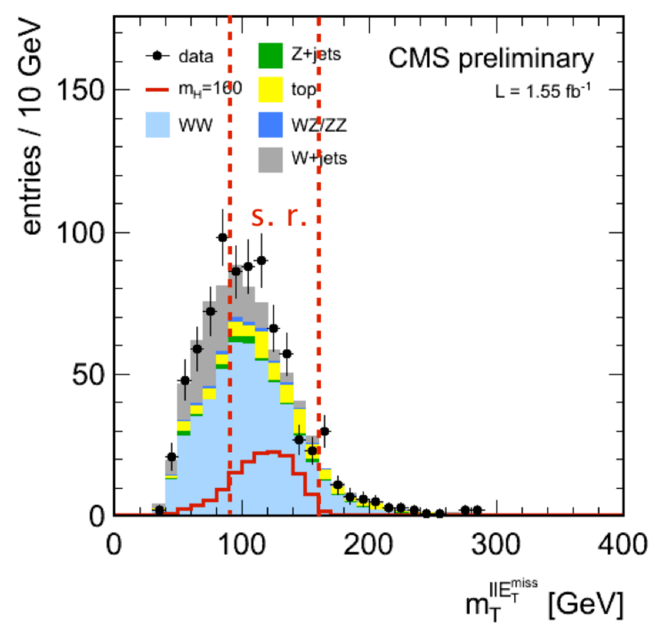
For large Higgs masses, the overlap between signal and EWK WW is large → WW background from MC



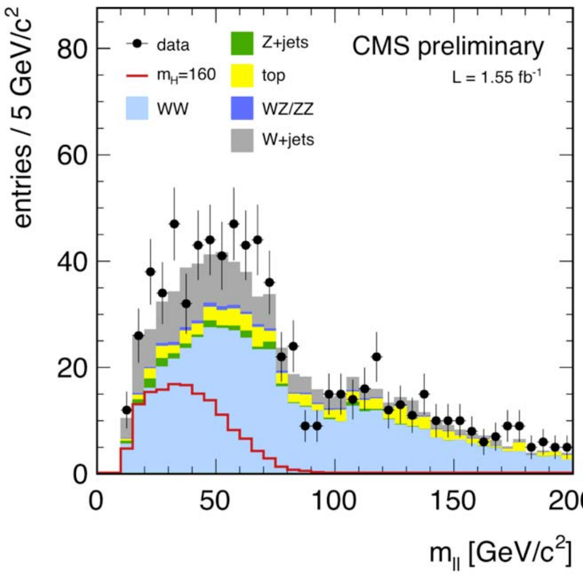
WW background further reduced with additional kinematic cuts

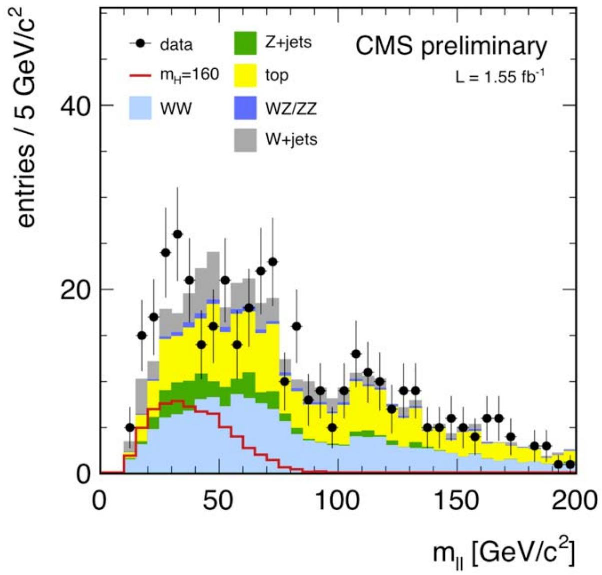


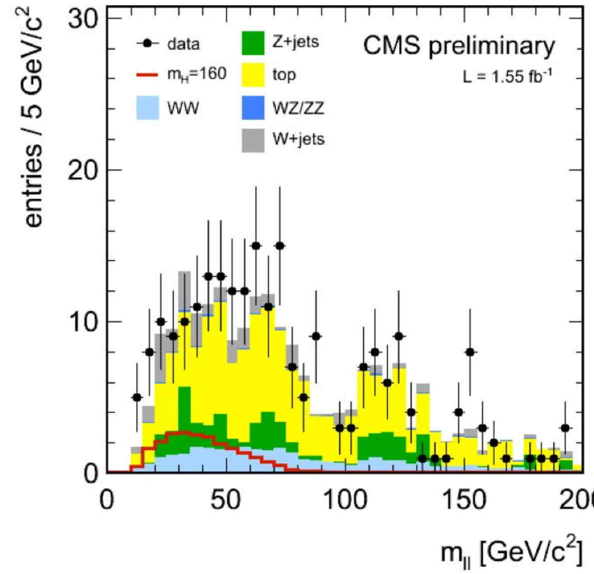




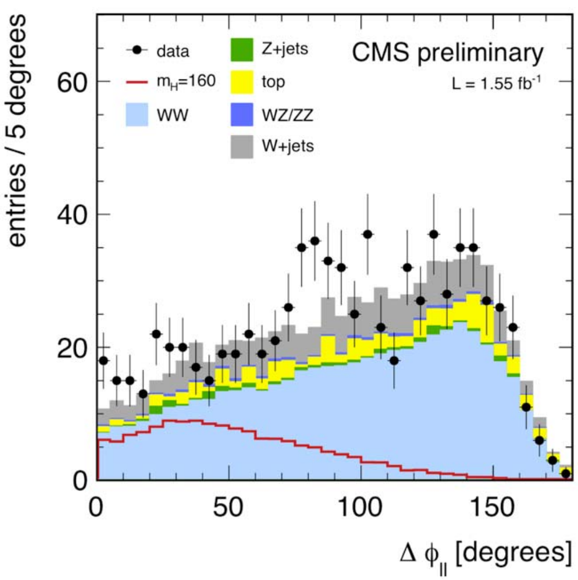
Kinematic Distributions

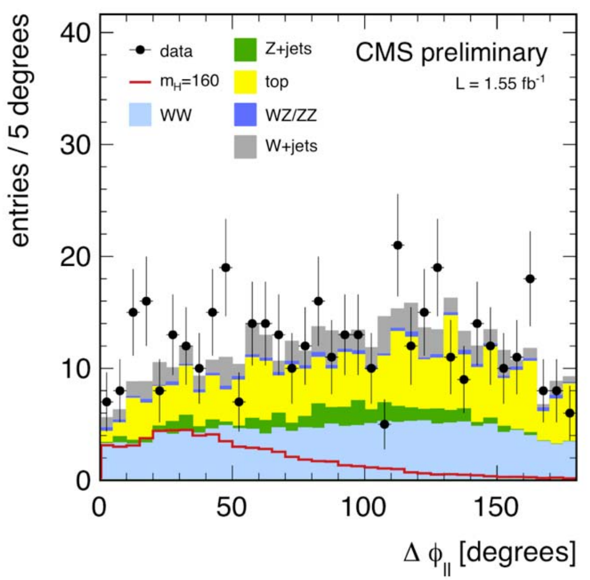


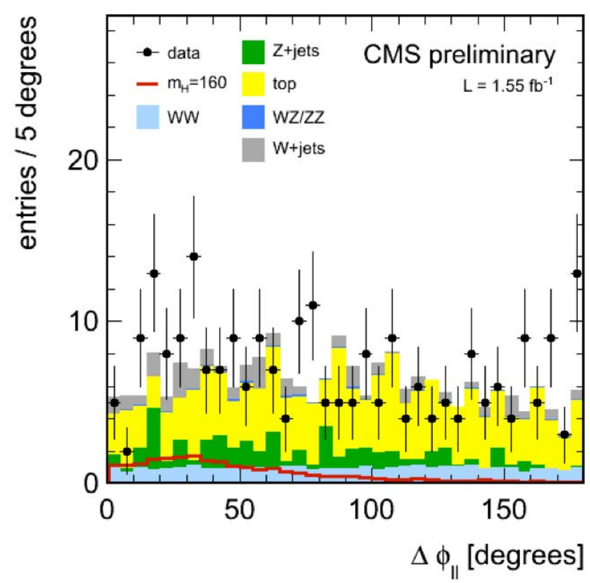




0-jet bin **1-jet bin** **2-jet bin**







Making Limits

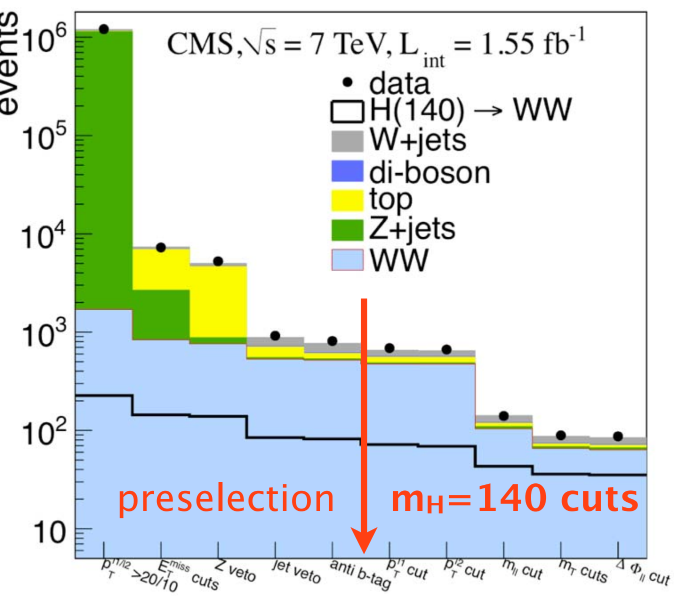
Systematics

Source	$H \to W^+W^-$	$qq \to W^+W^-$	$gg \to W^+W^-$	non-Z resonant WZ/ZZ	top	DY	W + jets	$V(W/Z) + \gamma$
Luminosity	4.5	—	—	4.5	—	—	—	4.5
Trigger efficiencies	1.5	1.5	1.5	1.5	—	—	—	1.5
Muon efficiency	1.5	1.5	1.5	1.5	—	—	—	1.5
Electron id efficiency	2.5	2.5	2.5	2.5	—	—	—	2.5
Momentum scale	1.5	1.5	1.5	1.5	—	—	—	1.5
E_T^{miss} resolution	2.0	2.0	2.0	2.0	2.0	3.0	—	1.0
Jet counting	7-20	—	5.5	5.5	—	—	—	5.5
Higgs cross section	5-15	—	—	—	—	—	—	—
WZ/ZZ cross section	—	—	—	3.0	—	—	—	—
$qq \to WW$ norm.	—	15	—	—	—	—	—	—
$gg \to WW$ norm.	—	—	50	—	—	—	—	—
W + jets norm.	—	—	—	—	—	—	36	—
top norm.	—	—	—	—	25	—	—	—
$Z/\gamma^* \to \ell^+\ell^-$ norm.	—	—	—	—	—	60	—	—
Monte Carlo statistics	1.0	1.0	1.0	4.0	6.0	20.0	20.0	10.0

- **Main signal systematics:**
  - Luminosity
  - Jet counting
  - Higgs cross section
- Main background systematics:
  - **intrinsic in the method**
  - **control sample definition**
  - **MC sample size**

Event yields after Higgs cuts

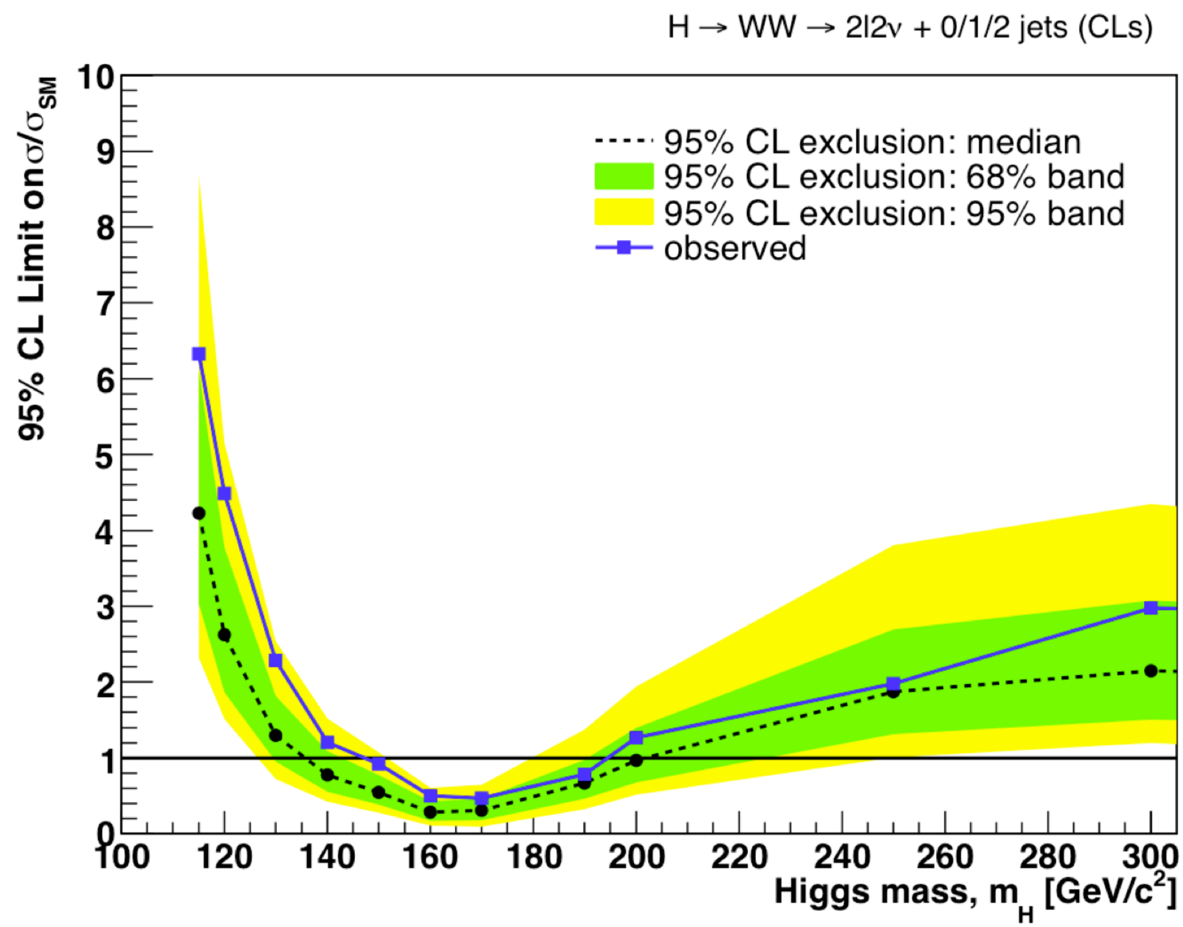
m_H=140 GeV/c^2 analysis



Process	0-j OF	0-j SF	1-j OF	1-j SF	2-j
qqWW	31.5 ± 5.5	29.1 ± 5.1	8.3 ± 3.1	5.8 ± 2.2	0.6 ± 0.2
ggWW	1.5 ± 0.8	1.3 ± 0.7	0.5 ± 0.3	0.3 ± 0.2	0.1 ± 0.1
VV	0.8 ± 0.1	0.5 ± 0.1	0.5 ± 0.1	0.3 ± 0.1	0.0 ± 0.0
Top	3.1 ± 1.1	1.4 ± 0.5	5.6 ± 1.2	3.2 ± 0.8	2.6 ± 1.5
Zjets	0.1 ± 0.0	3.1 ± 4.2	0.2 ± 0.1	1.2 ± 2.7	0.8 ± 0.6
Wjets	5.6 ± 2.3	5.3 ± 2.2	2.4 ± 1.1	1.5 ± 0.9	1.0 ± 0.6
$W\gamma$	1.5 ± 0.7	0.0 ± 0.0	0.0 ± 0.0	0.2 ± 0.2	0.0 ± 0.0
$Z\tau\tau$	0.0 ± 0.0	0.0 ± 0.0	0.2 ± 0.2	0.0 ± 0.0	0.2 ± 0.2
Tot. Bkg.	44.0 ± 6.2	40.6 ± 7.0	17.8 ± 3.5	12.6 ± 3.7	5.3 ± 1.7
Higgs	19.1 ± 4.3	16.1 ± 3.6	7.7 ± 2.6	5.3 ± 1.8	2.5 ± 0.3
Data	46	41	23	23	7

You can guess what the limit will be...

Results



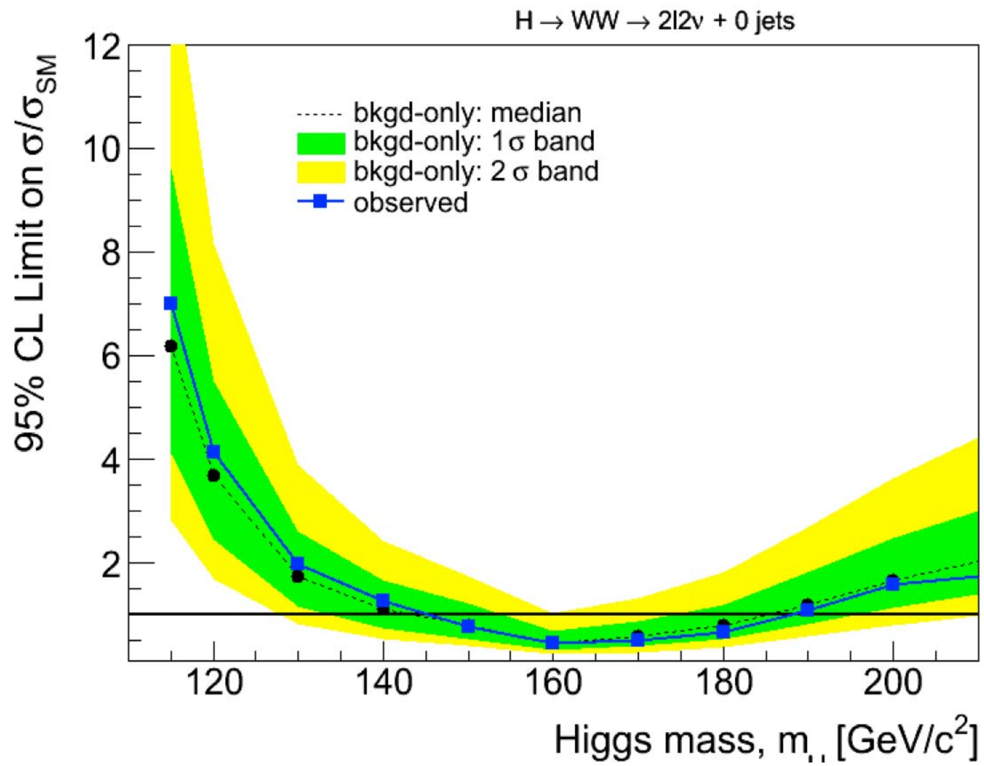
expected range: [136,200] GeV/c^2

observed 95% CL exclusion

for m_H in range [147,194] GeV/c^2

Outlook

- At the moment the sensitivity for low mass Higgs is still low
- But most of the interest is now in this region...
- In the next period we need to focus on how to further improve at low m_H
- – Add more data
- – Update MVA analysis for 1.5/fb and more
- – Reduce systematic errors, especially on the W+jets background, dominant in this region



Expected limits for background only case and **mean-observed limits in the presence of a Higgs** with **m_H = 120 GeV**

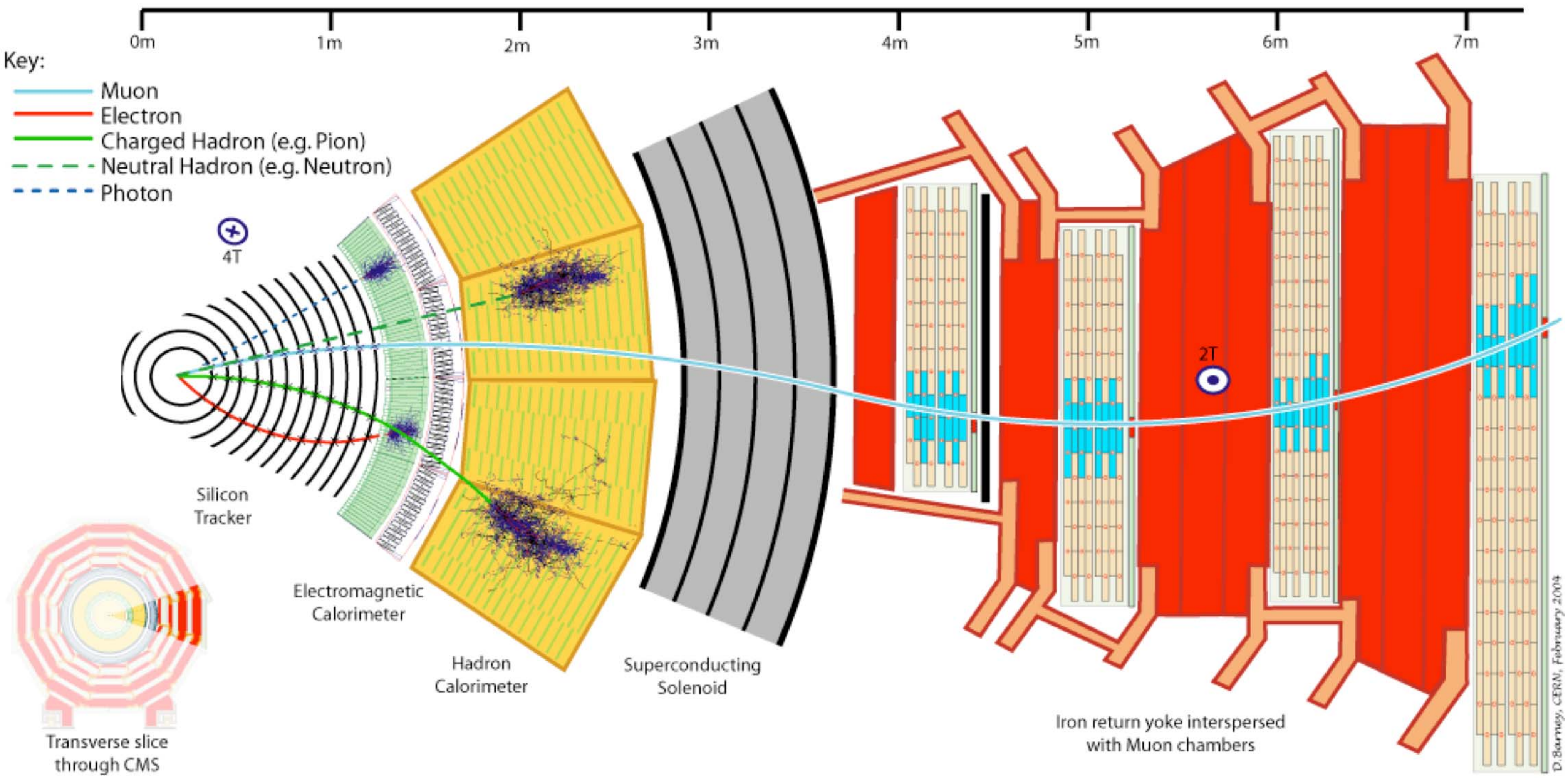
Plots made for L=1.1/fb and 0-jet bin only

Summary

- The CMS search for $H \rightarrow WW \rightarrow 2l2\nu$ with ~1.5/fb has been presented
- Crucial feature of this analysis is a precise background estimation in the high-intensity environment of LHC collisions
- Observed limits are [147,193] GeV/c^2
- The Higgs is still hiding... need more data to gain sensitivity in the low-mass region!

backup

CMS Detector



Analysis cuts

- Preselection:
  - leading (trailing) lepton p_T>20 (10) GeV
  - min-MET>20 (40) GeV for OF (SF) events
  - m_{ll}>12 GeV
  - $|m_{ll}-m_Z|$>15 GeV, $\Delta\phi_{ll,j}<165^\circ$ (SF only)
  - b-tag and soft muon veto
- 0/1 jet bins:

Table 1: Values of the selection requirements for several characteristic m_{H} masses.

m_{H}[GeV]	$p_{\mathrm{T}}^{\ell,\max}$ [GeV/c]	$p_{\mathrm{T}}^{\ell,\min}$ [GeV/c]	$m_{\ell\ell}$ [GeV/c^2]	$\Delta\phi_{\ell\ell}$ [dg.]	$m_{\mathrm{T}}^{\ell\ell E_{\mathrm{T}}^{\mathrm{miss}}}$ [GeV/c^2]
	$>$	$>$	$<$	$<$	[,]
130	25	10	45	90	[75,125]
150	27	25	50	90	[80,150]
160	30	25	50	60	[90,160]
180	36	25	60	70	[120,180]
200	40	25	90	100	[120,200]
300	70	25	200	175	[120,300]

- 2-jet bin:
  - $|\Delta\eta_{j1j2}|$>3.5
  - m_{j1j2}>450 GeV/c^2

preselection yields

	data	all bkg.	$qq \to W^+W^-$	$gg \to W^+W^-$	$t\bar{t} + tW$	$W + \gamma$
0-jet	811	771.2 ± 70.0	494.8 ± 44.6	23.8 ± 2.2	72.6 ± 18.8	12.3 ± 2.3
1-jet	435	427.6 ± 32.1	152.1 ± 13.8	8.2 ± 0.8	156.3 ± 19.8	3.4 ± 1.0
2-jet	252	235.4 ± 22.3	33.2 ± 3.1	1.5 ± 0.1	131.7 ± 16.8	1.6 ± 0.7

	WZ/ZZ not in $Z/\gamma^* \to \ell^+\ell^-$	$WZ + ZZ + Z/\gamma^* \to \ell^+\ell^-$	$Z/\gamma^* \to \tau^+\tau^-$	W + jets
0-jet	12.0 ± 1.3	15.2 ± 5.1	1.9 ± 0.6	138.5 ± 50.2
1-jet	10.1 ± 1.1	16.3 ± 4.3	14.9 ± 2.2	56.3 ± 20.6
2-jet	2.2 ± 0.3	28.3 ± 11.5	4.3 ± 0.9	22.6 ± 8.6

Triggers

- Two sets of triggers are used, analysis and utility triggers

- **Analysis Triggers**
  - Mainly use of dilepton triggers
  - Additional efficiency from unprescaled single lepton triggers
    - high pT thresholds and tight ID cuts
  - total efficiency for signal events >97%

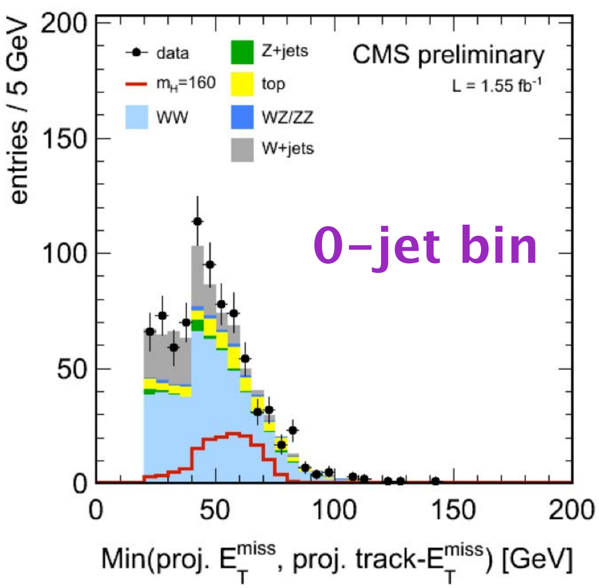
- **Utility Triggers**
  - analysis triggers require ID on both legs, cannot be used for utility measurements
  - utility triggers for efficiency
  - utility triggers for fake rate measurement

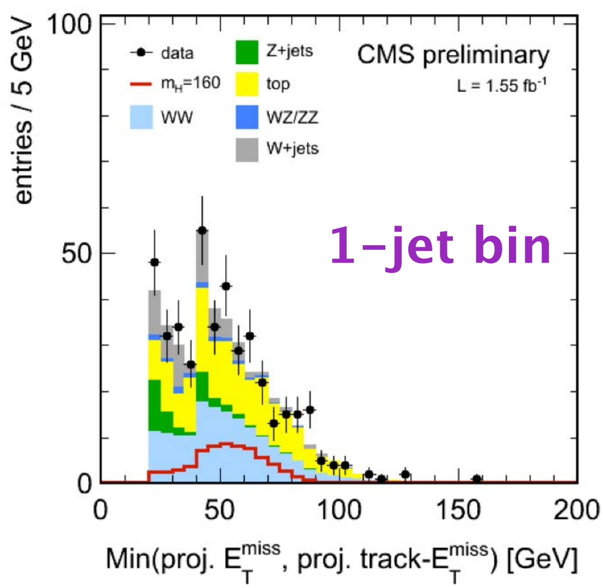
Missing Energy

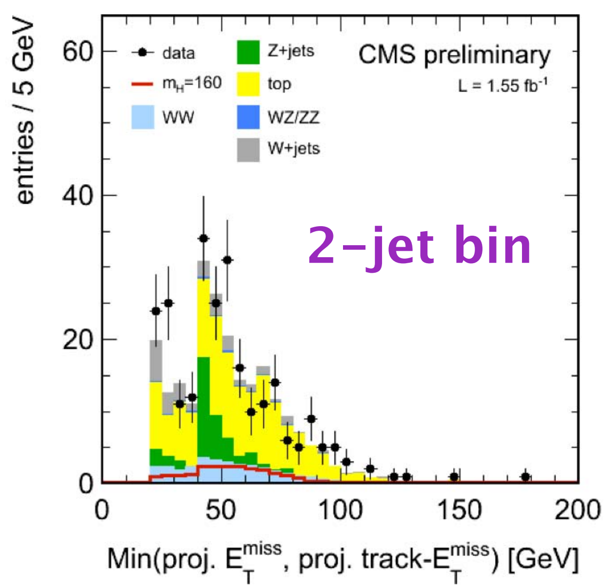
$$\Delta\phi_{min} = min(\Delta\phi(\ell_1, E_T^{miss}), \Delta\phi(\ell_2, E_T^{miss}))$$

$$\text{projected } E_T^{miss} = \begin{cases} E_T^{miss} & \text{if } \Delta\phi_{min} > \frac{\pi}{2}, \\ E_T^{miss}\sin(\Delta\phi_{min}) & \text{if } \Delta\phi_{min} < \frac{\pi}{2} \end{cases}$$

$$\text{min-MET} = min(proj.\ E_T^{miss}, proj.\ track\text{-}E_T^{miss})$$

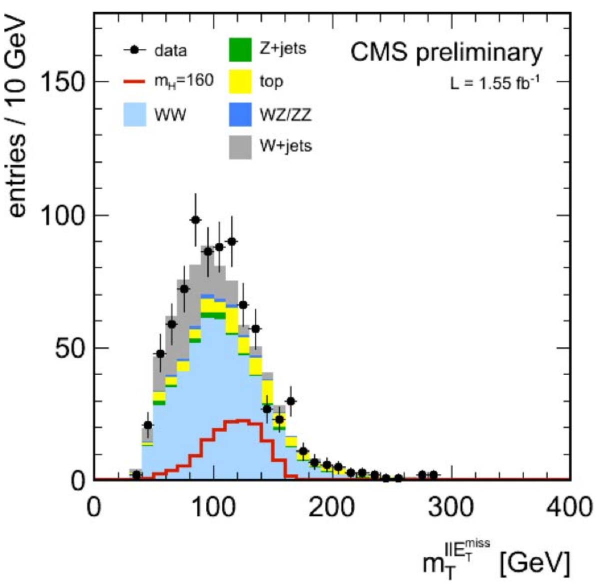


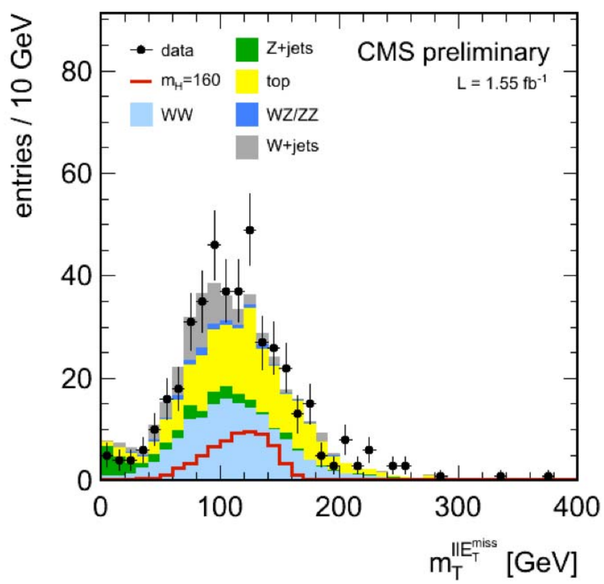


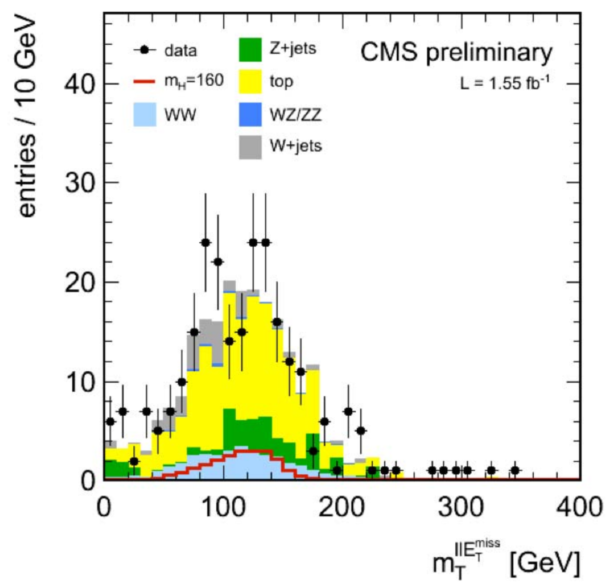


Transverse Mass

$$m_T^{\ell\ell E_{\mathrm{T}}^{\mathrm{miss}}} = \sqrt{2p_{\mathrm{T}}^{ll} E_{\mathrm{T}}^{\mathrm{miss}} (1 - cos(\Delta\phi_{\ell\ell - E_{\mathrm{T}}^{\mathrm{miss}}}))}$$







b-tagging

- Top events are identified using two methods: soft muon and b-tagging.

- The soft muon method tags events with an extra muon with $p_T > 3$ GeV and passing looser ID requirements

- Standard b-tagging is applied on jets
  - track counting algorithm
  - consider not-counted jets with $p_T > 7$ GeV
    - improve top rejection in the 0-jet bin
  - efficiency measured on data well in agreement with MC
  - ~50% efficiency on top, ~2% mistag rate

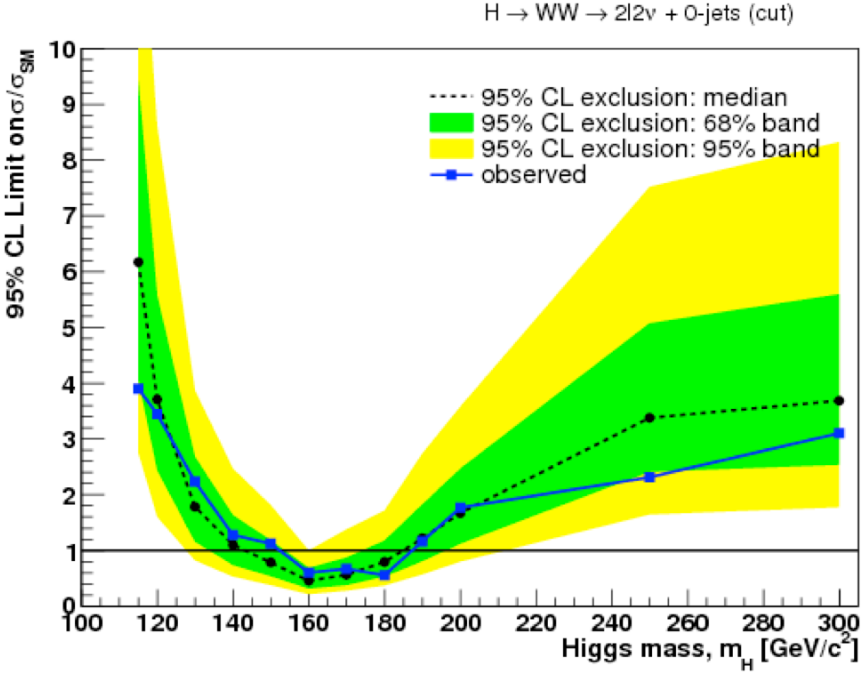
W+jets systematics

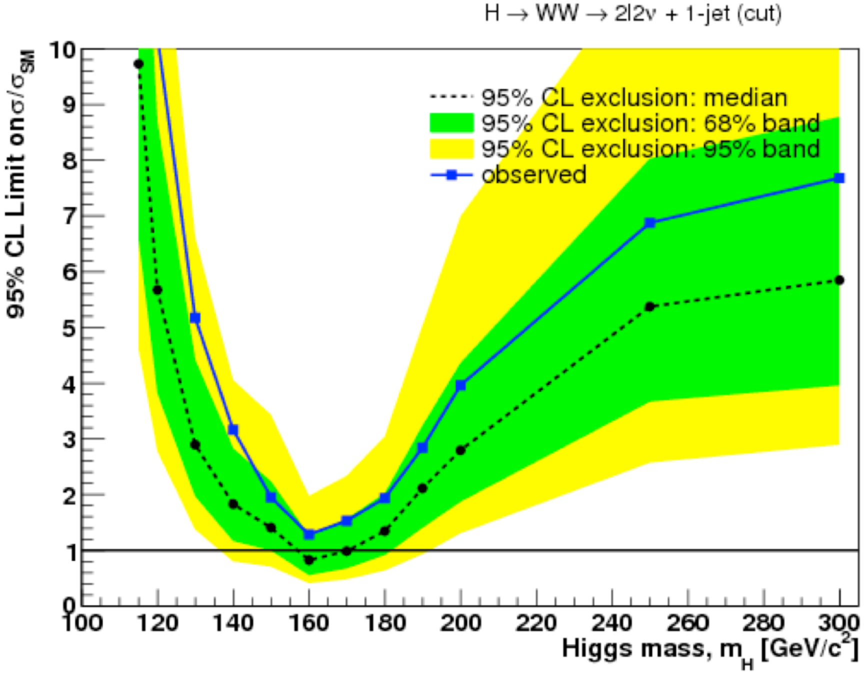
- Two sources:
  - different fake composition in measurement and application sample
  - different event kinematics
- Evaluation of the uncertainty:
  - MC closure test: measurement from QCD sample and application on W+jet sample
  - variation of secondary (recoiling) jet p_T threshold
- Overall systematic uncertainty is 36%
  - Needs to be improved does not scale with lumi (i.e. requires R&D...)

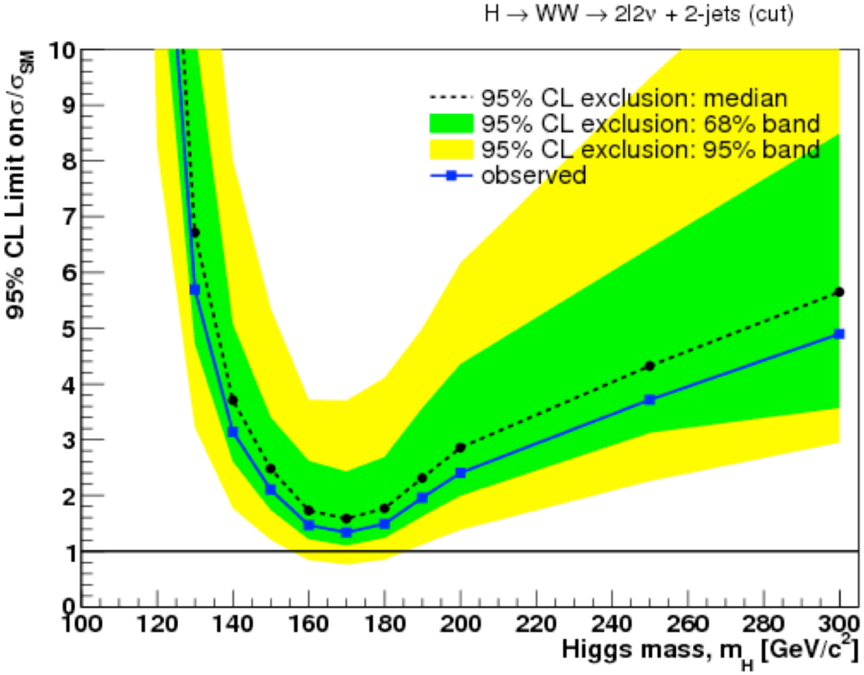
A deeper insight into the limit plots

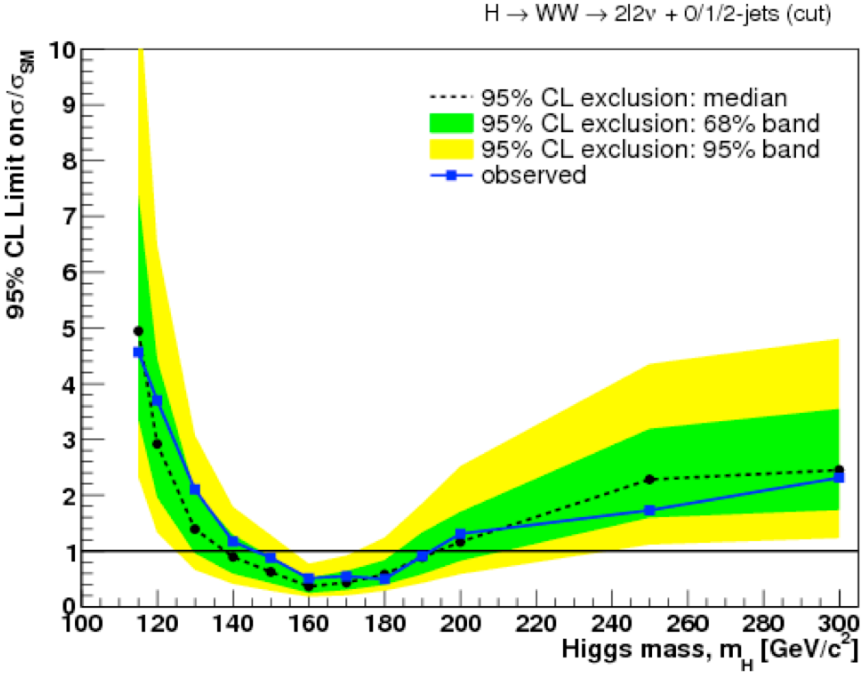
Shown for 1.1/fb, bayesian approach

Limits split by jet bins

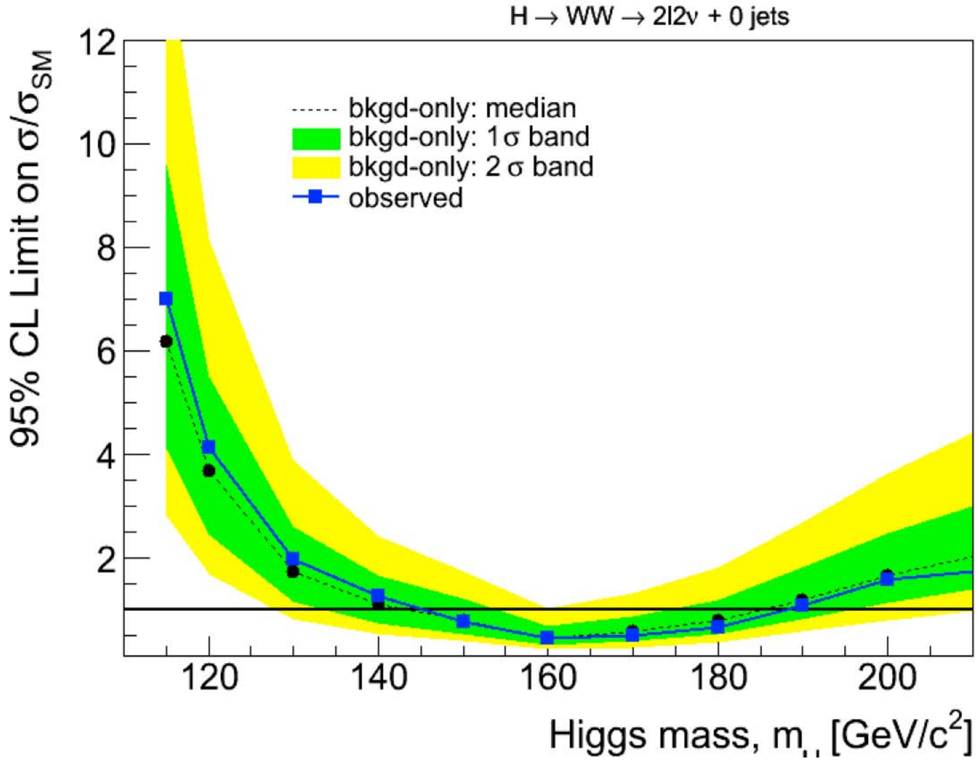


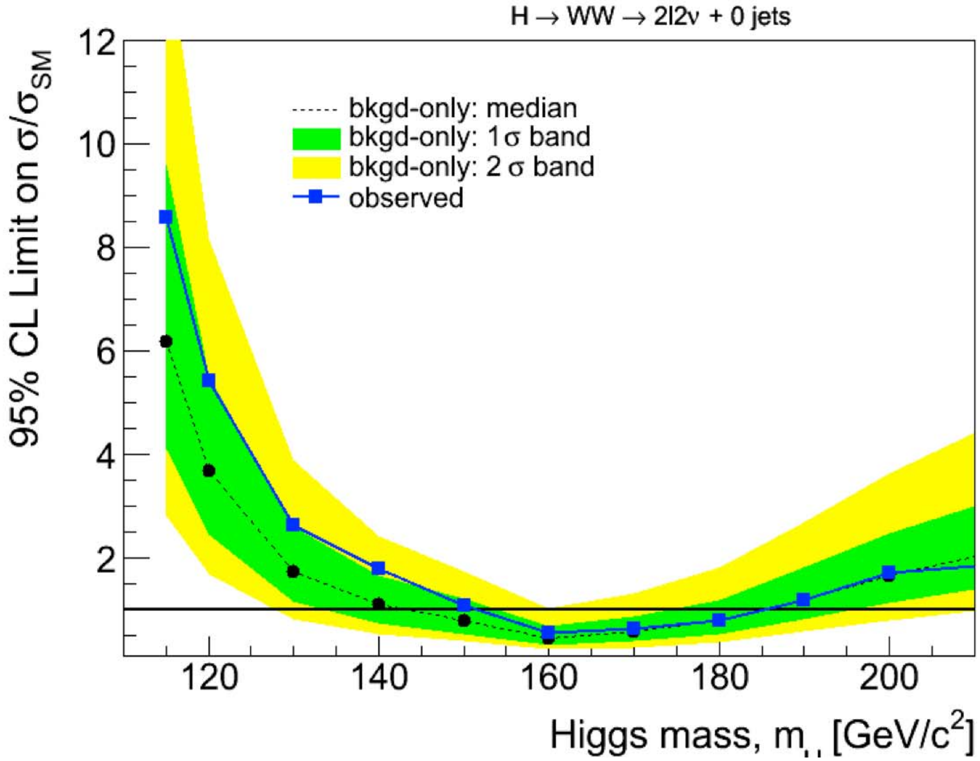






Higgs or not Higgs?





Expected limits for background only case and **mean-observed limits in the presence of a Higgs** with **m_H = 120 GeV (left)** and **m_H = 130 GeV (right)**

Plots made for L=1.1/fb and 0-jet bin only

Brazilian flag